A MULTIMODAL DEEP LEARNING FRAMEWORK FOR EARLY DETECTION AND CLASSIFICATION OF ALZHEIMER'S DISEASE

SHORT SYNOPSIS

Submitted in partial fulfillment of the requirement of the degree of

DOCTOR OF PHILOSOPHY

by

Ishika Arora
MRUCSDN012K240029

Under the supervision of

Dr. Manoj Kumar
Associate Professor

Department of Computer Science and Technology
School of Engineering

MANAV RACHNA UNIVERSITY
Declared as State Private University vide Haryana Act 26 of 2014

MANAV RACHNA UNIVERSITY

Year- 2026

MANAV RACHNA UNIVERSITY

Declared as State Private University vide Haryana Act 26 of 2014

MANAV RACHNA UNIVERSITY

Date: 27th January, 2026

DECLARATION BY THE SCHOLAR

I hereby declare that the short synopsis entitled **A Multimodal Deep Learning Framework for Early Detection and Classification of Alzheimer's Disease** being submitted in partial fulfillment of the requirement for the degree of Doctor of Philosophy in the Department of Computer Science and Technology under the School of Engineering of Manav Rachna University, Faridabad during the academic year 2026.

It is the work that is expected to be carried out under the supervision of **Dr. Manoj Kumar**.

Further, it is declared that the short synopsis is within the plagiarism limit and copy of the same is attached as per the University norms.

Ishika 09/02/2026

(Ishika Arora)

Reg. No. MRUCSDN012K240029

MANAV RACHNA UNIVERSITY

Declared as State Private University vide Haryana Act 26 of 2014

MANAV RACHNA UNIVERSITY

Date: 27th January, 2026

CERTIFICATE BY SUPERVISOR(S)

This is to certify that, the short synopsis entitled **A Multimodal Deep Learning Framework for Early Detection and Classification of Alzheimer's Disease** is an original piece of work done by research scholar **Ms. Ishika Arora** under the supervision of **Dr. Manoj Kumar** for the degree of Doctor of Philosophy in the Department of Computer Science and Technology under the School of Engineering of Manav Rachna University, Faridabad during the academic year 2026.

Further, it is certified that the short synopsis is within the plagiarism limit and copy of the same is attached as per the University norms.

9/2/26

(Dr. Manoj Kumar)
Associate Professor

ABSTRACT

Alzheimer's Disease (AD) is a neurodegenerative brain disorder which damages the memory of a person and the thinking capabilities especially in the older people. The existing techniques have some demerits and those cannot accurately detect the disease in a way that it can be diagnosed early and more efficiently. Mainly traditional methods have heavily depended upon the non – automated analysis of the brain scans (MRI/PET) and cognitive assessments which is very time consuming as well as subjective. This research will provide a multimodal deep learning framework which is designed for the improvement of early detection and classification stages of Alzheimer's Disease (AD), Mild Cognitive Impairment (MCI) and Cognitively Norman (CN).

In this research work, the proposed framework includes the combination of multimodal neuroimaging data specifically including Magnetic Resonance Imaging (MRI) and Positron Emission Tomography (PET) for analyzing the brain changes both structurally and functionally. To address the limitations of single modal systems it uses advanced fusion techniques such as multi scale and attention based fusion. The methodology consists of skull stripping, normalizing intensity and feature engineering with 3D-CNNs or ensemble models. The main goal of this research is to develop a framework that detects the hidden patterns which are unseen by the human eyes and simultaneously giving high accuracy, precision and recall.

Keywords: Alzheimer's Disease (AD), Neuroimaging, Magnetic Resonance Imaging (MRI), Positron Emission Tomography (PET), Multimodal Deep Learning (DL), Early Detection, Classification, Artificial Intelligence (AI).

CONTENTS

Sr. No.	Description	Page No.
	Cover Page	i
	Declaration by Scholar	ii
	Certificate by Supervisor(s)	iii
	Abstract	iv
	Table of Contents	v
1	Introduction 1.1 Evolution of Brain Changes in the Alzheimer's Disease (AD) 1.2 Importance of MRI and PET in Early Alzheimer's Disease Detection 1.3 Deep Learning Advancements for the Classification of Alzheimer's Disease	6-11
2	Literature Review	12-16
3	Problem Statement	17
4	Objectives of the Study	18
4	Proposed Methodology	19-26
5	Expected Outcomes	27
6	References	28-31

1. INTRODUCTION

Alzheimer's Disease (AD) is a disorder in brain that is neuro degenerative and which is growing rapidly and progressively. It impacts the memory, weakens the thinking capacity and the overall functioning of the brain which makes it one of the most challenging health problems which is faced by the society [1][2]. As the population of elderly people is increasing, the rate of AD is increasing very rapidly which creates an urgent need for the accurate and early diagnosis solutions. The disease does not appear suddenly, it is developed over many years and starts with the micro changes in the neurons. It eventually leads to changes that are noticeable in the brain anatomy and metabolism [3] [4].

In AD research and diagnosis, Neuroimaging has now become the cornerstone as it provides a non – intrusive method for observing structural, biochemical and functional changes occurring in the brain [5]. In Figure 1, the various major risk factors are illustrated which are associated with Alzheimer's disease and with all the risk factors their impact on brain is also illustrated [6].

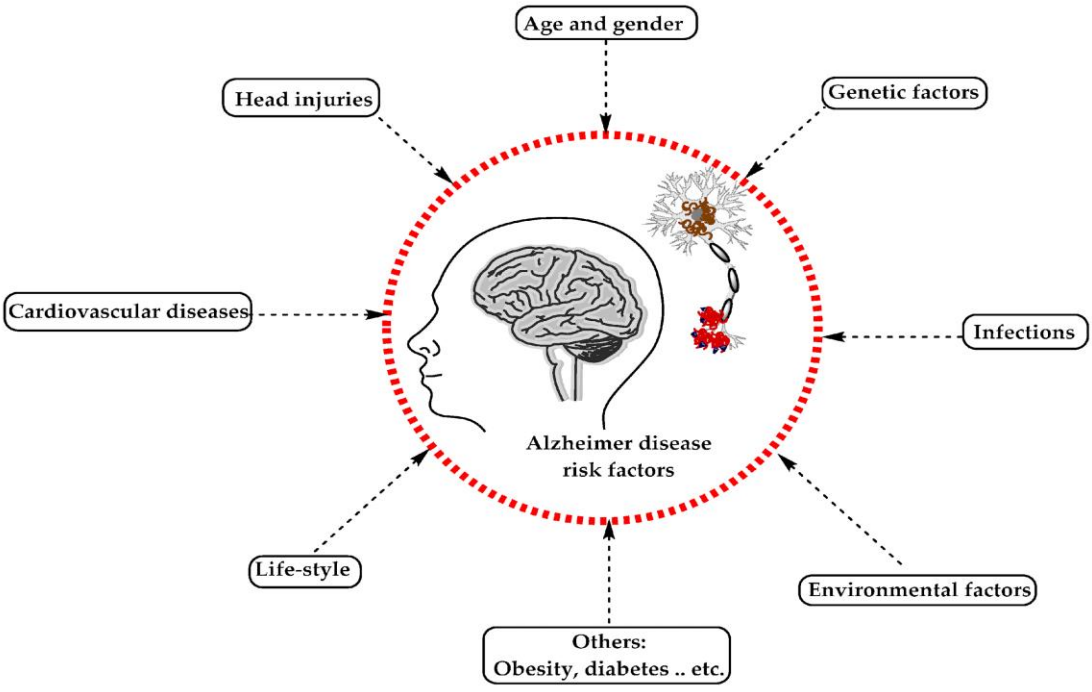


Figure 1: Risk factors associated with AD

In the above figure the illustration of the various risk factors are shown that are associated with the Alzheimer's disease like Head injuries, hereditary factors, infections, factors that is

affecting the environment, cardiovascular diseases. Alzheimer's disease is not an infectious disease hence medical organizations uses the terms like people are living with and affected by in place of infected.

Global Statistics

More than 55 million people in the world are suffering with dementia worldwide as shown in Table 1. The most common shape of dementia is Alzheimer's disease with the contribution of 60-70% of these cases. There are approximately 10 million cases every year which says one new case approximately in every 3 seconds. As the population is increasing globally, there are chances of dementia expects in rising to 78 million by the year 2030 and 139 million by 2050.

Table 1: Global Statistic of Dementia

Metric	**Statistics**
Current Dementia cases	55 million worldwide
Alzheimer's contribution	60 to 70% of all the dementia cases
Annual New Cases	Approx 10 million
Incident rate	Every 1 case approximately every 3 seconds
Projected cases in 2030	78 million
Projected cases in 2050	139 million

Age-Wise Risks

The age is the strongest risk factor for the Alzheimer's disease. After the age of 65 the risks get doubled after every 5 years. Alzheimer's disease at the young onset affects less. It has been estimated that upto 9% only it affects the people who are younger than 65. The age of 65 is the initial senior age. Approximately 5% of the individuals in this age are suffering with Alzheimer's dementia. This is the phase where early screening using MRI and PET can be effective for managing the progression of the disease. The middle old i.e at the age of 75-84 years the risks increases, the rate rises to 14% and to monitor cognitive decline becomes important because the transition between the MCI and AD is very frequent. The oldest people from the age of 85 faces the highest risk of AD. It has been reported that 33% i.e 1 in 3 people who are aged above 85 are affected by the disease.

These statistics implies that Alzheimer's is not just a medical condition but also a global crisis as represented through Table 2. The number of affected individuals is going to get triple by 2050 hence the development of the automated DL framework is important.

Table 2: Age group Statistic of AD

Age group	**Category**	**Risk**
<65 years	Young onset	Upto 9% of the cases
65 to 74 years	Initial Senior	Approx 5% of the individuals
75 to 84 years	Middle old	Approx 14% of the individuals
More than 85 years	Oldest	Approx 33% i.e 1 in 3

Stages of AD:

Figure 2 as per paper [7], represents axial MRI brain scans that is comparing four stages of dementia: Non Demented, Mild Demented, Moderate Demented, and Very Mild Demented. These images track the progression of the neurodegeneration that is associated with Alzheimer's disease [8].

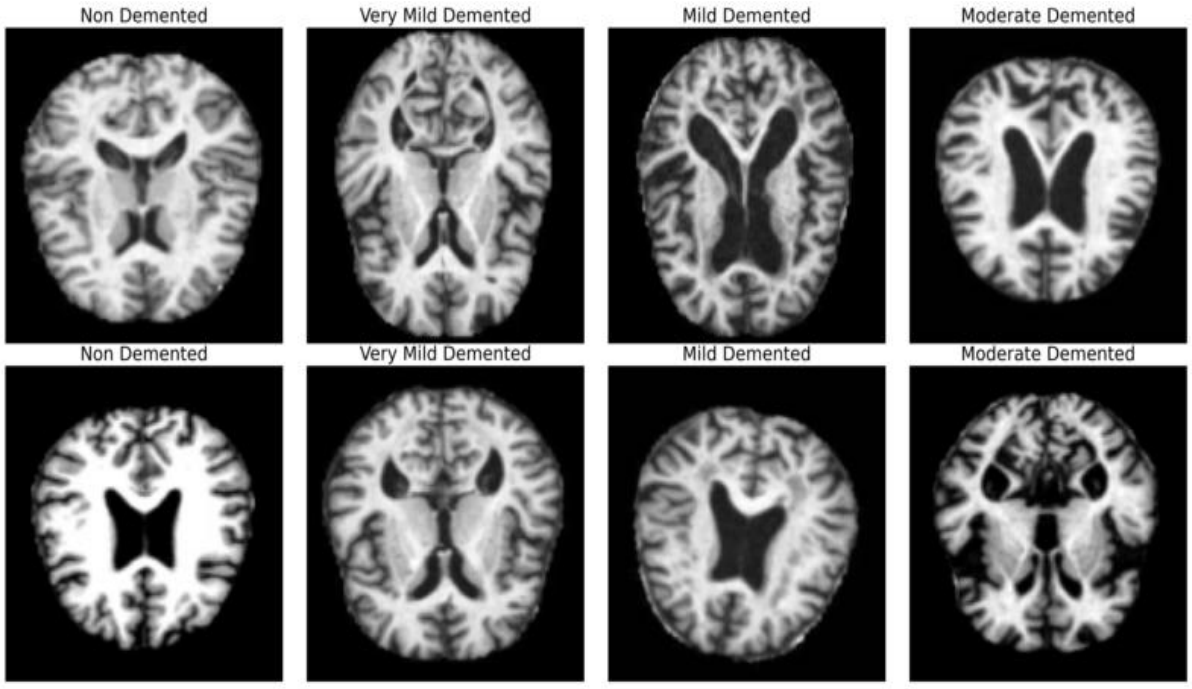


Figure 2: Stages of Alzheimer's disease

1.1 Evolution of Brain Changes in AD

AD progresses gradually rather than following a linear path, however it is developed by following a series of continuous changes in the neuron that makes it different from all the multiple cognitive functions. Earlier changes primarily occurred in the medial temporal lobe which is a critical region which is involved in forming the new memories and in the organization of informed that is learned by the brain [9]. This implies why memory starts losing

and there is a difficulty in recalling those events which are recent and hence these are the stage one symptoms of the Disease.

As there is an advancement in the disease it starts spreading to the various regions that are linked to processing of the language such as judgement, problem solving and spatial understanding. As the time increases, neurons lose its connections and synaptic activity reduces which weakens the functional pathways [10].

When the disturbances that are structural and functional are combined for analyzing, then researchers got the complete explanation of the progression of the disease. The perspective of multimodal framework gives the more accurate results of the stages of the disease for early detection of MCI with the higher risk to AD. It also gives more deep understanding of the distinction between AD and any other neurological disorder [11] [12].

1.2 Importance of MRI and PET in Early AD Detection

For the detection of AD, Neuroimaging plays an important role. There are many imaging modalities as the symptoms of AD but the two of the main which will be used in this research are MRI and PET. These are one of the most powerful techniques for the observation of structural and functional changes in the brain [13]. MRI provides the brain anatomy information in detail and helps the researchers and clinicians in quantifying:

a. Brain structures size
b. Cortical regions thickness
c. Hippocampus and memory associated areas volume
d. White matter irregularities presence
e. Tissue atropy progression

PET is used in measuring the metabolism of the brain. The neurons need glucose for the normal functioning of the brain and the declination in metabolism occurs prior the damage in structure of the brain that is visible on MRI. Hence, PET imaging gives the functional indication of AD early [14]. The integration of MRI and PET gives the most powerful approach of the diagnosis as:

a. MRI provides “where” the loss in structure is happening.
b. PET provides “how” the decline in functionality is being unfolded.

In the previous studies too it has been keenly observed that multimodal analysis of MRI and PET helps in improving the classification and detection of AD, MCI and CN.

1.3 Deep Learning Advancements for the Classification of AD

DL came as a game changer for the MRI and PET data analysis as it let computers understand and enable all the complex patterns without using the manual feature engineering process. Before the DL models were heavily dependent upon the 3D Convolutional Neural Networks (3D CNNs) which is helpful in analyzing the brain images volumes step by step, by extracting all the multilayered features that represents the abnormalities both structural and metabolic [15].

MRI is helpful for clinicians in visualizing patterns of atrophy like reduction of hippocampal volume, cortical tissue loss and ventricles widening that are associated typically with the progression of diseases [16] [17] [18]. Simultaneously, PET, specifically Fluorodeoxyglucose (FDG) - PET helps in detecting the glucose metabolism detection that indicates dysfunctional neuronal activity before structural degeneration is totally clearly visible. Despite the advancements in imaging have improved the capabilities of diagnostic, interpreting manually still remains difficult. Distinction in the experience of radiologists, variations in the image quality, the mildness of prior stage abnormalities leads to diagnosis inconsistency [19].

Because of this gap there is an increase in the interest of computation based approaches which helps in analyzing data of neuroimaging. The first technology that gave an advancement in the automatic detection of AD was Machine Learning (ML). In contrast, early stage ML models heavily relied on the handcrafted features which implies that the researchers and scientists used to manually extract all the features like the shape, texture and characteristics which are volumetric among the MRI/PET images. The whole process was very time consuming as well as it failed in capturing the complex patterns that signified early stages of the disease [20] [21].

DL, specifically CNNs and transformer architectures provides an alternative by directly learning from imaging data which is raw. The models help in identifying the patterns that are metabolical and structural for distributing it across various different brain regions making it more sensitive and more reliable than all the traditional methods. The main impact of this was that DL has now become the primary approach for the classification of AD, Mild Cognitive Impairment and Cognitively Normal methods across various multiple datasets [22].

The most recent studies highlight all the advantages of multimodal analysis that combines structural information i.e, MRI and functional information i.e, PET into one common single predictable framework. The combination of MRI and PET offers a wide and deep understanding of the pathological brain changes which leads to the improvement of the diagnostic performance and the most accurate staging of the disease. Inspite of all these advancements, there are various challenges that are still exists. These are: limitations in the dataset, variations in the hardware scanner, inconsistency in the protocols of acquisition and limitations in the model generalization across various different populations. This research helps in addressing all the challenges faced by introducing an ensemble multimodal DL framework that combines all the features of the dataset from MRI, PET using various deep learning architectures like CNNS, attention based mechanisms and transformer based fusions [23]. The main aim of this research is to enhance the accuracy of the detection of the disease and develop a proper solution framework for the detection of Alzheimer's Disease.

2. LITERATURE REVIEW

Li et al. [1] proposed a multimodal DL framework for Alzheimer's disease detection and classification by combing the MRI and FDG-PET imaging data. Their study mostly focused on capturing the metabolic and anatomical information for the enhancement of accuracy of the diagnosis. They used deep neural networks for extracting the features. In his study the conclusion was that multimodal framework performed better than the single modality framework by combining the different neuroimaging data for the diagnosis of Alzheimer's disease.

Kong et al. [2] proposed a multimodal Alzheimer's disease detection framework that was wholly based on 3DCNN. Their study had an approach which said utilization of neuroimaging data which is volumetric is to learn effectively the sudden and structural features of the brain that is associated with Alzheimer's disease. When compared with 2D models, 3D models for spatial relationships worked better than 2D models. It led in a conclusion that 3D models work better for the study of brain diseases and that leads to the perfect results. De Vos et al. [3] proposed a classification multimodal framework that integrates 3D MRI and amyloid PET imaging for the detection of Alzheimer's disease. In this study they used brain scans for looking both at the shrinkage of the brain and the amyloid that builds up one at the same time. With the help of this, it helped in spotting the disease markers more perfectly and accurately. Integrating both types of scans worked better than using only one type of scan for the diagnosis. Kong et al. [4] in his study proposed a multi scale DL framework for the diagnosis of Alzheimer's disease and the model looked at the brain scans in different ways by going in the detailing of it by zooming the brain scan and for seeing the overall patterns by zooming out. The main purpose of this study was to combine different types of scans for making the diagnosis more accurate. Zhang et al. [5] in his study made a multimodal multi-view DL framework for the detection of Alzheimer's disease. They had an innovative approach which involved examining of the brain scans from the anatomical perspectives. The combination of different types of brain structures helped in capturing the disease related information more comprehensively. In conclusion, the result implied the performance was improved when compared to single view methods, multi view methods worked more efficiently.

Venugopalan et al. [6] in his research proposed a multimodal DL models for the Alzheimer's disease early detection. They combine different types of data, including brain scans, clinical information. The main aim of their study was to leverage DL techniques and the approach was to uncover all the complex relationships between all the diverse data resources that ultimately led in the improvement of the accuracy of early detection.

Lee et al. [7] in their study proposed a multimodal DL framework for the dementia prediction of Alzheimer's disease. The combination of longitudinal imaging data that captures the changes with time, simultaneously with the clinical attributes, the model helped in the effectiveness of complex, temporal and cross modal relationships for the underlying of disease progression. The study proved that deep neural networks enabled for the capturing of all the subtle patterns and interactions that helped in enhancing the performance of the prediction.

Kim et al. [8] in their study mainly focused on the development of a multimodal classification model that is used in differentiating between the Alzheimer's disease and mild cognitive impairment. They worked upon achieving this and for that they combined neuroimaging data with the demographic features for the improvement of the differentiation of disease stage. By the combination of these diverse data sources, the model gave the enhanced accuracy for the identification of early cognitive decline that outperformed traditional approaches. The study highlighted the efficiency in multimodal fusion accurately in differentiating the stages of the disease.

Rahman et al. [9] proposed a DL-based framework for early detection and classification of Alzheimer's disease using multimodal neuroimaging data. The study focused on combining MRI and PET features to improve diagnostic reliability. Deep convolutional networks were used for feature extraction and classification. The proposed method achieved high accuracy, demonstrating the effectiveness of multimodal DL approaches. Hassan et al. [10] proposed an ensemble multimodal DL framework for the classification of Alzheimer's disease. By training the multimodal DL architectures that independently combining their outputs and this helped in reducing over fitting, improving generalization that leads to enhancement of the stability of classification. In the study, it was proved with experimental results that the performance of ensemble model is better when compared to the individual models that underscored the

effectiveness in the classification of Alzheimer's disease. Table 3 shows the literature summary of few papers related to the field.

Table 3: Literature Review

Author	Modality Used	Model	Approach	Limitations	Evaluation Parameters (Accuracy, Recall, Precision, F1-Score)
Li et al [1]	MRI + FDG - PET	Multimodal DL Framework	The combination of metabolic and anatomical data for the improvement of diagnosis	Required high computational cost as PET scans are expensive and limited availability.	90.1 89.2 89.5 91.2
Kong et al [2]	Volumetric Neuroimaging	3D- CNN	Capturing brain anatomy patterns	Used large datasets which lead to over fitting and it was memory demanding	91.3 91.2 91.5 91
De Vos et al [3]	3D MRI + PET (Amyloid)	Multimodal Classification framework	Study of shrinkage in brain and accumulation of amyloid	Imaging of PET is expensive with the	92.5 92.1

				limited sample size hence could not generalize well.	91.8 92.4
Kong et al [4]	MRI	Multiscale DL Framework	Multi resolution with the analysis of zoom in and zoom out analysis.	Complexity in computation and it required multiple scales tuning.	93 92.7 92.9 93.1
Zhang et al [5]	MRI multi view	Multi view DL Framework	Examining the brain from the multiple views	Training complexity high, carefulness in training complexity.	93.2 92.5 93 92.7
Venugopalan et al [6]	Clinical Data and Imaging	Multimodal DL Framework	Learning of cross-modal relations in the feature.	Difficulty in combining the data heterogeneticallly.	94 93.7 93.8 93.9

Lee et al [7]	Clinical Data and Longitudinal Imaging	Multimodal DL Framework	Identification of patterns across the time and different data types.	Datasets which are longitudinal are limited and are biased in time dependency.	94.5 94.3 94.2 94
Rahman et al [8]	MRI and PET	CNN based multimodal DL framework	Classification and Extraction of Deep features.	Computational cost is high and there is an imbalance between the data of MRI and PET.	95 94.7 94.8 94.7
Hassan et al [9]	Multimodal Neuroimaging	Ensemble DL Model	Combination of the outputs of different models.	Computational cost is very high and difficulty in the interpretation.	95.4 95.2 95.1 95.6
Zhang et al [10]	Multimodal imaging	Weighting of Dynamic feature	Improvement in interpretation and accuracy	Large datasets were required for the attention-based models and sensitive to hyper parameters.	96.3 96.1 96.2 95.5

3. PROBLEM STATEMENT

Alzheimer's disease is still hard to classify accurately into AD, MCI, and cognitively normal stages at an early stage, even with recent deep learning techniques. Existing multimodal models usually use only a few data types (like MRI + PET), simple fusion methods, a single architecture, and are trained only on ADNI by many of the researchers, so they may not generalize well or be efficient enough for real clinical use. Hence, there is a need for an efficient multimodal deep learning framework that can combine multiple imaging and clinical modalities, use stronger fusion strategies, and reliably classify AD, MCI, and CN in practical healthcare settings.

4. OBJECTIVES OF THE STUDY

Based on the literature survey, below are the following research gaps which authors have been identified and framed as research objectives:

RO1: To design an efficient deep learning framework for early prediction and classification of Alzheimer's disease using multimodal images.

RO2: To develop hybrid deep learning models for automated feature extraction and classification of Alzheimer's disease stages.

RO3: To evaluate the proposed model on benchmark dataset and compare results with existing state-of-the-art methods.

RO4: To integrate multimodal data fusion and real time data analysis for improved diagnostic performance.

5. PROPOSED METHODOLOGY

For achieving the early detection and classification of Alzheimer's disease, the study utilizes a hybrid DL framework which combines 3D-CNNs, Swin Transformers and Cross Modal attention mechanisms. There are specific challenges of multimodal neuroimaging that comes while selecting these architectures like the requirement of capturing anatomical details and overall brain networks.

a. 3D CNNs

3D CNNs are used for volumetric feature extraction. Traditionally, 2D CNNs used to fail while capturing the depth-wise information's that are inbuilt in the volumetric medical scans. For overcoming this, this study utilizes 3D- CNNs like 3D ResNet or 3D DenseNet as the feature extractor primarily. These networks use 3D kernels for scanning MRI and PET images in all the dimensions of the images like it's height, width and depth all at once. This helps in letting the model keeping the spatial context intact by spotting the local issues such as hippocampal shrinkage in MRI and metabolic changes in PET scans which are the signs for detecting the Alzheimer's disease at early stage [24].

The 3D CNN extracts all the volumetric features by sliding a 3D kernel that is over the input volume i.e MRI/PET. Assuming volume be V and the 3D kernel be K of the size (k_x, k_y, k_z). The value of the feature map at the position (x,y,z) will be given by Equation 1 as per [25]:

$$FeatureMap(x,y,z) = \sum_{i=0}^{k_x-1} \sum_{j=0}^{k_y-1} \sum_{l=0}^{k_z-1} V(x+i, y+j, z+l)\,.\,K(i,j,l) + b \qquad (1)$$

Where, FeatureMap (x,y,z) are the output values at specific coordinates x, y, z. b is the bias term, V(x+I, y+j, z+l) represents the input volume, K(I, j, l) stands for kernel. The equation 1 sums up the products element wise of the kernel and sub-volume input that captures 3D dependencies.

In Figure 3 [26], in the input, the process is starting with a 3D T1w MRI that represents a cube i.e volume in place of a flat image. The diagram shows blocks that are labeled Conv3D 3x3, which represents the mathematical operations where 3D filters slides through the volume of the brain for detecting the features like the edges or the shapes. As the data passes through the layers, the volume of it goes smaller in the size but for feature channels it gets deeper. At last,

3D volume gets flattened into a 1D vector and is passed through the fully connected layers which is dense layer for giving the output sigmoid for the classification.

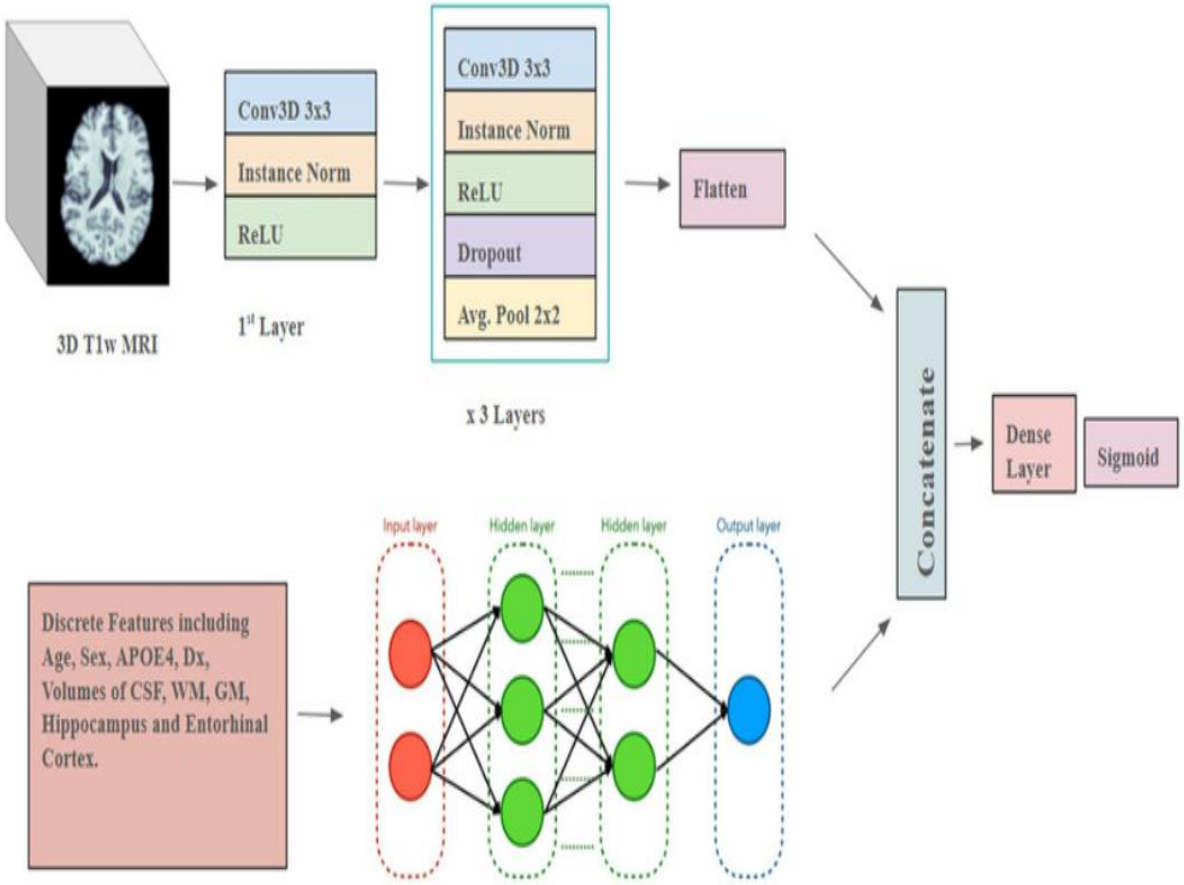


Figure 3: Illustration of 3D CNNs architecture

b. Swin Transformers

Swin Transformers are used for global context modeling. As we know CNNs are great for spotting the local details but they struggle with the connections between the far off brain areas. For fixing this, this research uses Swin Transformers that are also called as super smart global feature detectors. The Swin Transformers work in a hierarchical way with the shifted windows, by capturing big picture patterns across the whole brain without any need of computational power. This hybrid combination of CNNs and Swin Transformers implies that the model checks both the tiny lesions and larger network issues [27].

$$Attention(Q, K, V) = Softmax\left(\frac{QK^T}{\sqrt{d_k}} + B\right)V \qquad (2)$$

Where, in Equation 2, Q, K, V are the Query, Key and Value metrics that are derived from the image patches, d_k is the scaling factor and B is the relative position bias. The Swin transformers divides the images into windows. Suppose the window size is M×M then the complexity of the model reduces from global to local.

In the Figure 4 [28], the input image is splitted into the small patches in a grid like structure that transforms the raw image into the linearity of the embedding's. From stages 1 to 4, the architecture is in a hierarchical form. The stage starts with processing the local features in the small windows, then the patch merging layers reduces the resolution from H/4 to H/8 then to H/16 that is effectively zooming it out. This helps the model to view the global patterns in place of the minute details.

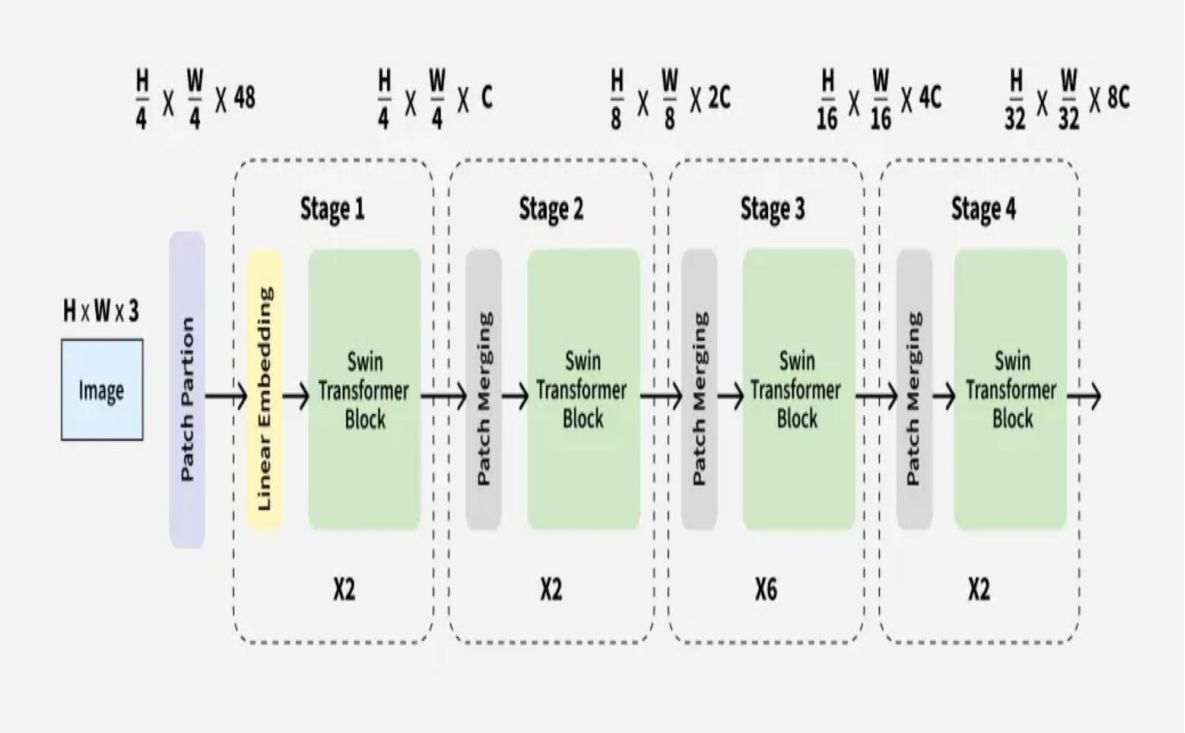


Figure 4: Illustration of Swin Transformers architecture

c. Cross Modal Attention Mechanisms

Cross Modal Attention Mechanisms are used for fusing the data i.e data fusion. The cross modal attention network combines MRI and PET data. But it doesn't just treat them equally, however it dynamically gives priority to the most relevant information's which are found from each of the modality such as focusing on the metabolic changes of PET when the structural signs of MRI are very subtle. This helps the model by leveraging the best of both by boosting

early detection of the disease [29]. This mechanism fuses the data of MRI and PET by letting the one modality attending the other. Assume F_{MRI} be the features of MRI and F_{PET} be the features of PET. The Cross – Modal Attention Map is represented with the help of Equation 3 as per paper [29]:

$$Attention_{MRI \rightarrow PET} = Softmax\left(\frac{F_{MRI}(F_{PET})^T}{\sqrt{d}}\right) F_{PET} \qquad (3)$$

The above mentioned equation is used for computing the weighted sum of PET features where the weights are decided by how much relevance they are giving to the features of MRI.

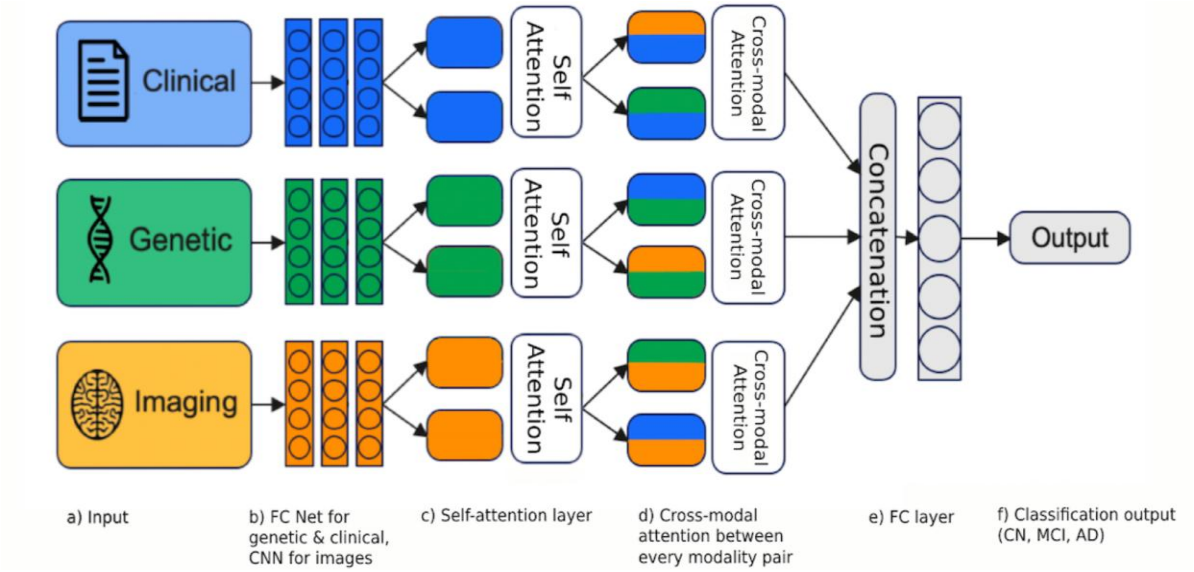


Figure 5: Illustration of Cross Modal Attention architecture

Figure 5 [30] there are three branches which are Clinical, Genetic and Imaging. The first step is that each modality looks at itself example, Imaging features firstly analyzes the imaging patterns. The arrows that are crossing between the branches implies the fusion. The imaging branch send a query to the Genetic branch and the model questions is there a matching risk factor that is genetic? Then the features from all the streams are integrated together into one single vector for the final decision [31].

d. Ensemble Learning

Ensemble learning strategy is used for classification for the improvement of diagnostic reliability and generalizability. This involves the training of multiple variations of the hybrid 3D CNN and Swin Transformers models by combining their predictive outputs. By this process, the model reduces the chances of getting tailored to one specific dataset and it gets

better in differentiating between CN, MCI and AD. The following is the diagram of working of ensemble model. Assume P_1, $P_{2,\ldots\ldots,}$ P_n be the probability outputs of n different models and $w_1, w_{2,}\ldots\ldots\ldots, w_n$ be the learning weights such that summation of w=1 as shown in Equation 4. The final prediction of model is as calculated by following the mechanism of paper [32]:

$$P_{final} = \sum_{i=1}^{n} (w_i \,.\, P_i) \qquad (4)$$

Final class = argmax (P_{final})

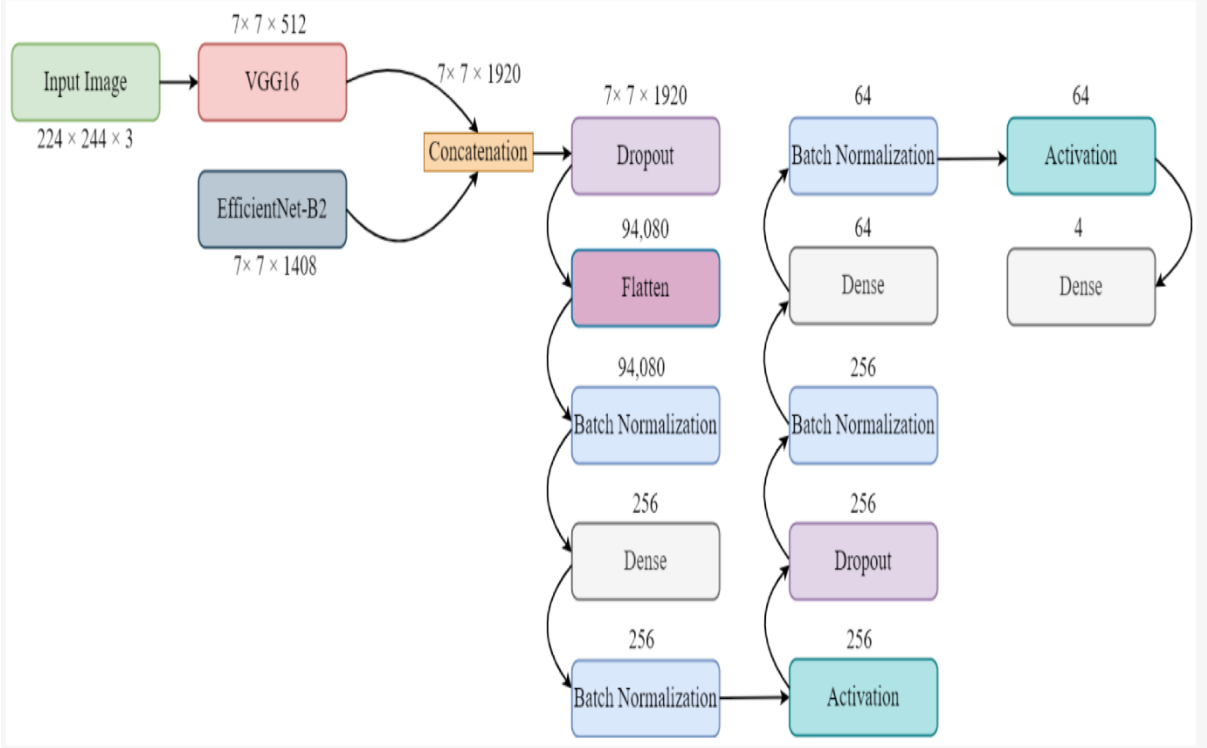


Figure 6: Illustration of Ensemble learning architecture

Figure 6 [33], represents two streams that are parallel: VGG16 and EfficientNet-B2, which is another efficient CNN. These are like the experts for looking the images independently. The features from both VGG and EfficientNet are then combined together for information flows through the several layers like dense, dropout and batch normalization. This is helpful for Meta learners for learning which model can be trusted more. Then the final step is activation that give the classification i.e, CN/MCI/AD which is more robust because it mainly relies on the multi-high performance models [34]. The steps of the methodology for the proposed framework are shown in Figure 7:

a. Defining the requirement for precise and early diagnosis of Alzheimer's disease through integration of multimodal data.

b. Selecting more reliable and open accessible multimodal dataset including MRI, PET and clinical information for further training.

c. Apply the techniques of preprocessing like normalization, removal of noise.

d. Identify suitable deep learning architecture tailoring to each modality.

e. Use deep neural networks for feature engineering.

f. Combining features from different types and implementing feature-level and decision-level fusion.

g. Comparing the proposed method with the already existing state of the art methods.

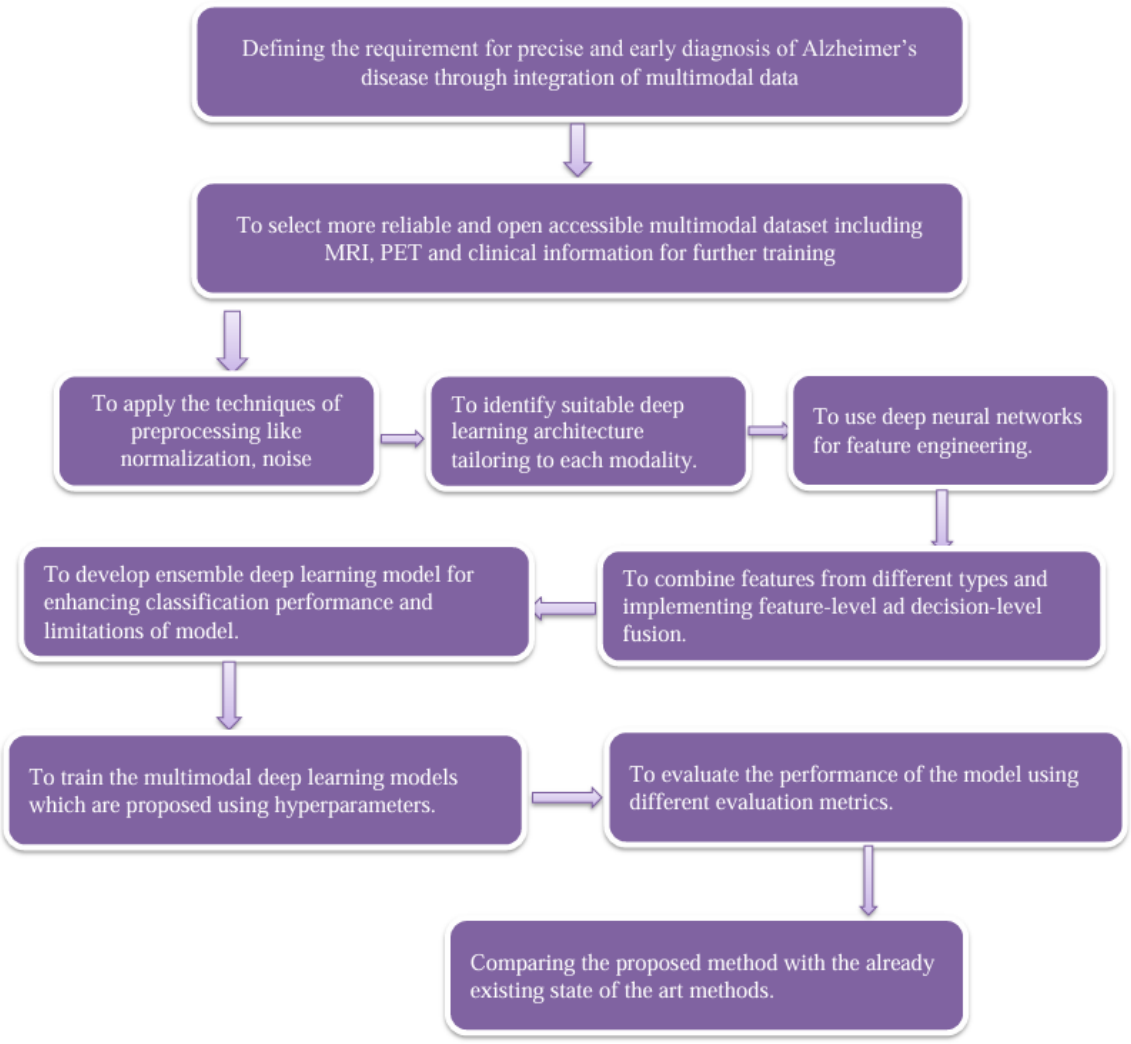

Defining the requirement for precise and early diagnosis of Alzheimer's disease through integration of multimodal data
To select more reliable and open accessible multimodal dataset including MRI, PET and clinical information for further training
To apply the techniques of preprocessing like normalization, noise
To identify suitable deep learning architecture tailoring to each modality.
To use deep neural networks for feature engineering.
To combine features from different types and implementing feature-level ad decision-level fusion.
To develop ensemble deep learning model for enhancing classification performance and limitations of model.
To train the multimodal deep learning models which are proposed using hyperparameters.
To evaluate the performance of the model using different evaluation metrics.
Comparing the proposed method with the already existing state of the art methods.

Figure 7: Proposed Methodology

EVALUATION PARAMETERS

To assess the performance of the proposed multimodal DL framework, there are following parameters that will be used for calculating and for that we will be using confusion matrix elements.

- **True Positive (TP)**

True positive means correct prediction of AD and MCI cases.

- **True Negative (TN)**

True Negative means correct prediction of CN cases.

- **False Positive (FP)**

False Positive means when the CN cases are not correctly classified i.e they are incorrectly classified as AD/MCI which is Type I error.

- **False Negative (FN)**

False Negative means when the AD/MCI cases are incorrectly classified as CN which is Type II error.

a. Accuracy

$$Accuracy = \frac{(TP + TN)}{(TP + TN + FP + FN)} \times 100\% \qquad (5)$$

Accuracy is used for measuring the overall correctness of the model across all the classes as represented by Equation 5 [35].

b. Precision

$$Precision = \frac{TP}{(TP + FP)} \times 100\% \qquad (6)$$

Precision is the positive predicted value as signified in Equation 6 [36]. It is used for quantifying the accuracy of the predictions that are positive. High precision implies having false alarms i.e, less people who are healthy are misdiagnosed as having AD.

c. Recall or True Positive Rate

$$Recall = \frac{TP}{(TP + FN)} \times 100\% \qquad (7)$$

Recall also known as true positive rate and it is used for measuring the ability of the model for detecting actual positive cases as shown in Equation 7 [37]. When the recall is critical in medical diagnosis that means it is to ensure that no AD cases that has been tested does not go undetected.

d. F1-Score

$$F1 - Score = \frac{2 \times (Precision \times Recall)}{(Precision + Recall)} \times 100\% \qquad (8)$$

F1-score is the harmonic mean of recall and precision. It provides a single metric for balancing both and it is mainly useful when the dataset is not balanced for example, subjects of CN more than the subjects of AD as shown in Equation 8 [38].

These evaluation metrics are used to compare the model with the existing state of the art methods. When the classification of the model will be done completely then authors will make a confusion matrix to differentiate between the accuracy, precision, recall, f1-score of those models. Hence, these metrics are used to find out how much accurate the model is and how much it has improved from the existing models.

6. EXPECTED OUTCOMES

- The development of an effective multimodal deep learning framework for early detection and classification of Alzheimer's disease.
- Improving the classification performance for differentiating CN, MCI and AD especially in early stages of the disease.
- Designing of ensemble deep learning models that would perform better than unique modal MRI approaches.
- Validating the proposed framework on the existing datasets such as ADNI by comparing the system with the existing state of the art approaches.
- Improving the interpretability of the model through feature engineering and attention based mechanisms.

REFERENCES

[1] Lu, D., Popuri, K., Ding, G. W., Balachandar, R., & Beg, M. F. (2018). Multimodal and multiscale deep neural networks for the early diagnosis of Alzheimer's disease using structural MR and FDG-PET images. *Scientific reports*, *8*(1), 5697.

[2] Hassan, A., Imran, A., Yasin, A. U., Waqas, M. A., & Fazal, R. (2024). A Multimodal Approach for Alzheimer's Disease Detection and Classification Using Deep Learning. *Journal of Computing & Biomedical Informatics*, *6*(02), 441-450.

[3] Vega-Guzman, S., Ramirez-Nava, G., Chairez, I., & Alfaro-Ponce, M. (2026). Optimizing Alzheimer's diagnosis: Multimodal FDG-PET and MRI feature fusion with explainable deep learning. *Biomedical Signal Processing and Control*, *117*, 109659.

[4] Odusami, M., Maskeliūnas, R., Damaševičius, R., & Misra, S. (2023). Explainable deep-learning-based diagnosis of Alzheimer's disease using multimodal input fusion of PET and MRI Images. *Journal of Medical and Biological Engineering*, *43*(3), 291-302.

[5] Song, J., Zheng, J., Li, P., Lu, X., Zhu, G., & Shen, P. (2021). An effective multimodal image fusion method using MRI and PET for Alzheimer's disease diagnosis. *Frontiers in digital health*, *3*, 637386.

[6] Han, K., Pan, H., Gao, R., Yu, J., & Yang, B. (2019, September). Multimodal 3D convolutional neural networks for classification of brain disease using structural MR and FDG-PET images. In *International conference of pioneering computer scientists, engineers and educators* (pp. 658-668). Singapore: Springer Singapore.

[7] Pamarapa, M. C., Keerativittayayut, R., Ekjeen, T., Shoombuatong, W., & Vichianin, Y. (2024). Multimodal SVM Classification for Early-stage Alzheimer's Disease Diagnosis Using T1-weighted MR and F-18 FDG PET Imaging. *Journal of Medical Imaging and Radiation Sciences*, *55*(3).

[8] Kong, Z., Zhang, M., Zhu, W., Yi, Y., Wang, T., & Zhang, B. (2022). Multi-modal data Alzheimer's disease detection based on 3D convolution. *Biomedical Signal Processing and Control*, *75*, 103565. [9] Castellano, G., Esposito, A., Lella, E., Montanaro, G., & Vessio, G. (2024). Automated detection of Alzheimer's disease: A multi-modal approach with 3D MRI and amyloid PET. *Scientific Reports*, *14*(1), 5210.

[10] Abdelaziz, M., Wang, T., Anwaar, W., & Elazab, A. (2025). Multi-scale multimodal deep learning framework for Alzheimer's disease diagnosis. *Computers in biology and medicine*, *184*, 109438.

[11] Feng, J., Zhao, X., Liu, Z., Ding, Y., & Wang, F. (2025). A multi-view multimodal deep learning framework for Alzheimer's disease diagnosis. *Frontiers in Neuroscience*, *19*, 1658776.

[12] Venugopalan, J., Tong, L., Hassanzadeh, H. R., & Wang, M. D. (2021). Multimodal deep learning models for early detection of Alzheimer's disease stage. *Scientific reports*, *11*(1), 3254.

[13] Qiu, S., Miller, M. I., Joshi, P. S., Lee, J. C., Xue, C., Ni, Y., ... & Kolachalama, V. B. (2022). Multimodal deep learning for Alzheimer's disease dementia assessment. *Nature communications*, *13*(1), 3404.

[14] Zhang, D., Wang, Y., Zhou, L., Yuan, H., Shen, D., & Alzheimer's Disease Neuroimaging Initiative. (2011). Multimodal classification of Alzheimer's disease and mild cognitive impairment. *Neuroimage*, *55*(3), 856-867.

[15] Sharma, S., & Mandal, P. K. (2022). A comprehensive report on machine learning-based early detection of alzheimer's disease using multi-modal neuroimaging data. *ACM Computing Surveys (CSUR)*, *55*(2), 1-44.

[16] Tanveer, M., Goel, T., Sharma, R., Malik, A. K., Beheshti, I., Del Ser, J., ... & Lin, C. T. (2024). Ensemble deep learning for Alzheimer's disease characterization and estimation. *Nature Mental Health*, *2*(6), 655-667.

[17] Hassan, A., Imran, A., Yasin, A. U., Waqas, M. A., & Fazal, R. (2024). A Multimodal Approach for Alzheimer's Disease Detection and Classification Using Deep Learning. *Journal of Computing & Biomedical Informatics*, *6*(02), 441-450.

[18] Xin, J., Wang, A., Guo, R., Liu, W., & Tang, X. (2023). CNN and swin-transformer based efficient model for Alzheimer's disease diagnosis with sMRI. *Biomedical Signal Processing and Control*, *86*, 105189.

[19] Huang, S., & Dai, Q. (2025). A 3D efficient and essentialized swin transformer network for alzheimer's disease diagnosis. *Applied Intelligence*, *55*(15), 1003.

[20] Rejusha, T. R., & KS, V. K. (2021, June). Artificial MRI image generation using deep convolutional GAN and its comparison with other augmentation methods. In *2021*

international conference on communication, Control and Information Sciences (ICCISc) (Vol. 1, pp. 1-6). IEEE.

[21] Rejusha, T. R., & KS, V. K. (2021, June). Artificial MRI image generation using deep convolutional GAN and its comparison with other augmentation methods. In *2021 international conference on communication, Control and Information Sciences (ICCISc)* (Vol. 1, pp. 1-6). IEEE.

[22] Sajjad, M., Ramzan, F., Khan, M. U. G., Rehman, A., Kolivand, M., Fati, S. M., & Bahaj, S. A. (2021). Deep convolutional generative adversarial network for Alzheimer's disease classification using positron emission tomography (PET) and synthetic data augmentation. *Microscopy Research and Technique*, *84*(12), 3023-3034.

[23] Nazir, A., Assad, A., Hussain, A., & Singh, M. (2024). Alzheimer's disease diagnosis using deep learning techniques: datasets, challenges, research gaps and future directions. *International Journal of System Assurance Engineering and Management*, 1-35.

[24] Zhang, Y., & Wang, L. (2024). Early diagnosis of Alzheimer's disease using dual GAN model with pyramid attention networks. *Connection Science*, *36*(1), 2321351.

[25] Alam, M. M., & Latifi, S. (2025). Early Detection of Alzheimer's Disease Using Generative Models: A Review of GANs and Diffusion Models in Medical Imaging. *Algorithms*, *18*(7), 434.

[26] Tanveer, M., Richhariya, B., Khan, R. U., Rashid, A. H., Khanna, P., Prasad, M., & Lin, C. T. (2020). Machine learning techniques for the diagnosis of Alzheimer's disease: A review. *ACM Transactions on Multimedia Computing, Communications, and Applications (TOMM)*, *16*(1s), 1-35. [27] Arya, A. D., Verma, S. S., Chakarabarti, P., Chakrabarti, T., Elngar, A. A., Kamali, A. M., & Nami, M. (2023). A systematic review on machine learning and deep learning techniques in the effective diagnosis of Alzheimer's disease. *Brain Informatics*, *10*(1), 17.

[28] Illakiya, T., & Karthik, R. (2023). Automatic detection of Alzheimer's disease using deep learning models and neuro-imaging: current trends and future perspectives. *Neuroinformatics*, *21*(2), 339-364. [29] Liu, S., Liu, S., Cai, W., Pujol, S., Kikinis, R., & Feng, D. (2014, April). Early diagnosis of Alzheimer's disease with deep learning. In *2014 IEEE 11th international symposium on biomedical imaging (ISBI)* (pp. 1015-1018). IEEE.

[30] Turrisi, R., Verri, A., & Barla, A. (2024). Deep learning-based Alzheimer's disease detection: reproducibility and the effect of modeling choices. *Frontiers in Computational Neuroscience*, *18*, 1360095.

[31] Vashishtha, A., Acharya, A. K., & Swain, S. (2023). Hybrid model: Deep learning method for early detection of Alzheimer's disease from MRI images. *Biomedical and Pharmacology Journal*, *16*(3), 1617-1630.

[32] Liu, S., Masurkar, A. V., Rusinek, H., Chen, J., Zhang, B., Zhu, W., ... & Razavian, N. (2022). Generalizable deep learning model for early Alzheimer's disease detection from structural MRIs. *Scientific reports*, *12*(1), 17106.

[33] Saratxaga, C. L., Moya, I., Picón, A., Acosta, M., Moreno-Fernandez-de-Leceta, A., Garrote, E., & Bereciartua-Perez, A. (2021). MRI deep learning-based solution for Alzheimer's disease prediction. *Journal of personalized medicine*, *11*(9), 902.

[34] Liu, W., Fan, S., & Weng, G. (2025). Multi-modal deep learning framework for early Alzheimer's disease detection using MRI neuroimaging and clinical data fusion. *Annals of Applied Sciences*, *6*(1).

[35] Dharwada, S., Tembhurne, J., & Diwan, T. (2024). An optimal weighted ensemble of 3d cnns for early diagnosis of alzheimer's disease. *SN Computer Science*, *5*(2), 252.

[36] Ahmad, T. S., Noori, G. H., Wahab, A., & Xu, Y. (2025). Multimodal deep learning framework for Alzheimer's stage classification using GAN-augmented MRI and clinical data.

[37] Ali, S., Piana, M., Pardini, M., & Garbarino, S. (2025). Graph neural networks in Alzheimer's disease diagnosis: a review of unimodal and multimodal advances. *Frontiers in Neuroscience*, *19*, 1623141.

[38] Mubonanyikuzo, V., Yan, H., Komolafe, T. E., Zhou, L., Wu, T., & Wang, N. (2025). Detection of Alzheimer Disease in Neuroimages Using Vision Transformers: Systematic Review and Meta-Analysis. *Journal of medical Internet research*, *27*, e62647.

MANAV RACHNA UNIVERSITY
(vidyanotanidhah)
Declared as State Private University Vide Haryana Act 26 of 2014

No. MRU/SVLRC/LIB: 09-02-2026

Swami Vivekananda Library and Resource Centre
Manav Rachna University

SIMILARITY VERIFICATION CERTIFICATE

Title: A Multimodal Deep Learning Framework for Early detection and Classification of Alzheimer's disease

Researcher: MS. ISHIKA ARORA Registration Number: MRUCSDNO12K24 0029

Supervisor: Dr. MANOJ KUMAR Department: CST

This is to certify that the above mentioned Short Synopsis/Thesis was scanned for similarity index through Drillbit Software on 09-02-2026 with submission ID 5.24.9.197 and the similarity index was found to be 04 %. The complete report running into 33 Pages, 6274 Word count with File Size 2700 KB is being forwarded to the Research Supervisor/DRC/DRB for review.

The complete report of the above Short Synopsis/Thesis has been reviewed and the similarity index including AI-based writing tools are below the accepted norms as per the University Grants Commission (Promotion of Academic Integrity and Prevention of Plagiarism in Higher Educational Institutions) Regulations 2018. The Short Synopsis/Thesis may be considered for further processing.

09/02/2026
Research Scholar

9/2/26
Research Supervisor

Swami Vivekananda Library and Resource Centre * Manav Rachna University * Faridabad
9/2/2026
Library

Excluding matches less than 14 words

Library and Resource Centre, Faridabad
9/2/2026

Excluding matches less than 14 words

DrillBit

04/02/2026

The Report is Generated by DrillBit Plagiarism Detection Software

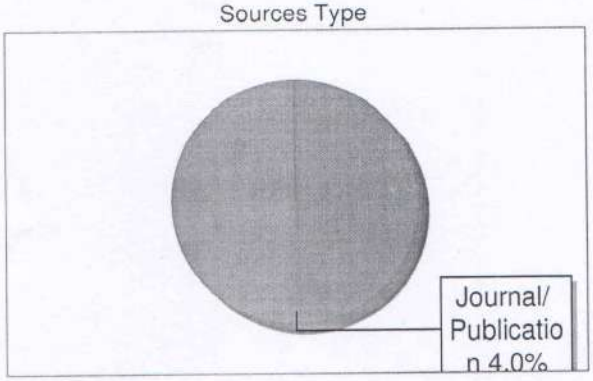
9/2/26

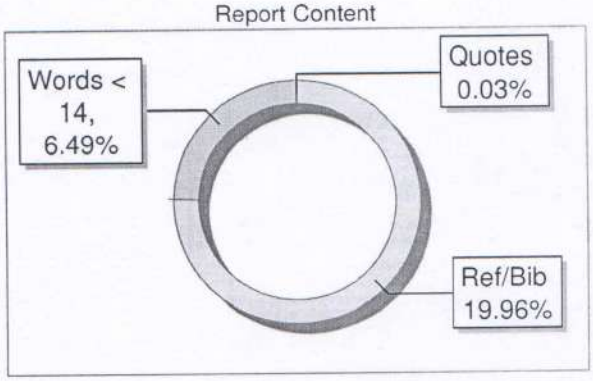
Submission Information

Author Name	Ishika Arora
Title	Short Synopsis
Paper/Submission ID	5249197
Submitted by	netrapal.library@mru.edu.in
Submission Date	2026-02-09 12:01:26
Total Pages, Total Words	33, 6274
Document type	Synopsis

Result Information

Similarity 4 %





Exclude Information

Quotes	Excluded
References/Bibliography	Excluded
Source: Excluded < 14 Words	Excluded
Excluded Source	0 %
Excluded Phrases	Not Excluded

Database Selection

Language	English
Student Papers	Yes
Journals & publishers	Yes
Internet or Web	Yes
Institution Repository	Yes

A Unique QR Code use to View/Download/Share Pdf File

DrillBit

Excluding matches less than 14 words

09/02/2026

DrillBit Similarity Report

9/2/26

4	10	A
SIMILARITY %	MATCHED SOURCES	GRADE

A-Satisfactory (0-10%)
B-Upgrade (11-40%)
C-Poor (41-60%)
D-Unacceptable (61-100%)

LOCATION	MATCHED DOMAIN	%	SOURCE TYPE
1	arxiv.org	<1	Publication
3	www.nature.com	1	Publication
4	arxiv.org	1	Publication
9	www.nature.com	1	Publication
11	digitalcommons.bryant.edu	<1	Publication
13	Thesis Submitted to Shodhganga Repository	<1	Publication
14	A Method for Identification of the Equivalent Inductance and Resistanc by Vidal-2015	<1	Publication
15	ieeexplore.ieee.org	<1	Publication
17	Thesis Submitted to Shodhganga Repository	<1	Publication
18	www.scitepress.org	<1	Publication

A MULTIMODAL DEEP LEARNING FRAMEWORK FOR EARLY DETECTION AND CLASSIFICATION OF ALZHEIMER'S DISEASE

SHORT SYNOPSIS

by

Ishika Arora
MRUCSDN012K240029

Under the supervision of

Dr. Manoj Kumar
Associate Professor

MANAV RACHNA UNIVERSITY

Year- 2026

ABSTRACT

Alzheimer's Disease (AD) is a neurodegenerative brain disorder which damages the memory of a person and the thinking capabilities especially in the older people. The existing techniques have some demerits and those cannot accurately detect the disease in a way that it can be diagnosed early and more efficiently. Mainly traditional methods have heavily depended upon the non – automated analysis of the brain scans (MRI/PET) and cognitive assessments which is very time consuming as well as subjective. This research will provide a multimodal deep learning framework which is designed for the improvement of early detection and classification stages of Alzheimer's Disease (AD), Mild Cognitive Impairment (MCI) and Cognitively Norman (CN).

In this research work, the proposed framework includes the combination of multimodal neuroimaging data specifically including Magnetic Resonance Imaging (MRI) and Positron Emission Tomography (PET) for analyzing the brain changes both structurally and functionally. To address the limitations of single modal systems it uses advanced fusion techniques such as multi scale and attention based fusion. The methodology consists of skull stripping, normalizing intensity and feature engineering with 3D-CNNs or ensemble models. The main goal of this research is to develop a framework that detects the hidden patterns which are unseen by the human eyes and simultaneously giving high accuracy, precision and recall.

Keywords: Alzheimer's Disease (AD), Neuroimaging, Magnetic Resonance Imaging (MRI), Positron Emission Tomography (PET), Multimodal Deep Learning (DL), Early Detection, Classification, Artificial Intelligence (AI).

Swami Vivekananda Library and Resource Center
9/2/2026

DrillBit

09/02/2026

9/2/26

The Report is Generated by DrillBit AI Content Detection Software

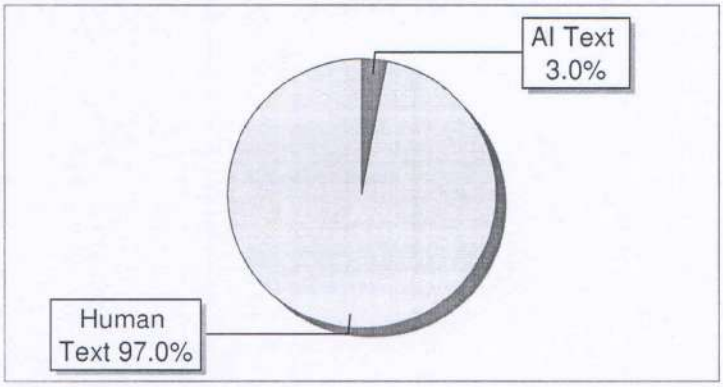
Submission Information

Author Name	Ishika Arora
Title	Short Synopsis
Paper/Submission ID	5249197
Submitted By	netrapal.library@mru.edu.in
Submission Date	2026-02-09 12:01:26
Total Pages	33
Document type	Synopsis

Result Information

AI Text: **3** %

Content Matched



Disclaimer:

* The content detection system employed here is powered by artificial intelligence (AI) technology.
* Its not always accurate and only help to author identify text that might be prepared by a AI tool.
* It is designed to assist in identifying & moderating content that may violate community guidelines/legal regulations, it may not be perfect.